AD-A063 864 MARYLAND UNIV COLLEGE PARK COMPUTER SCIENCE CENTER F/G 12/1

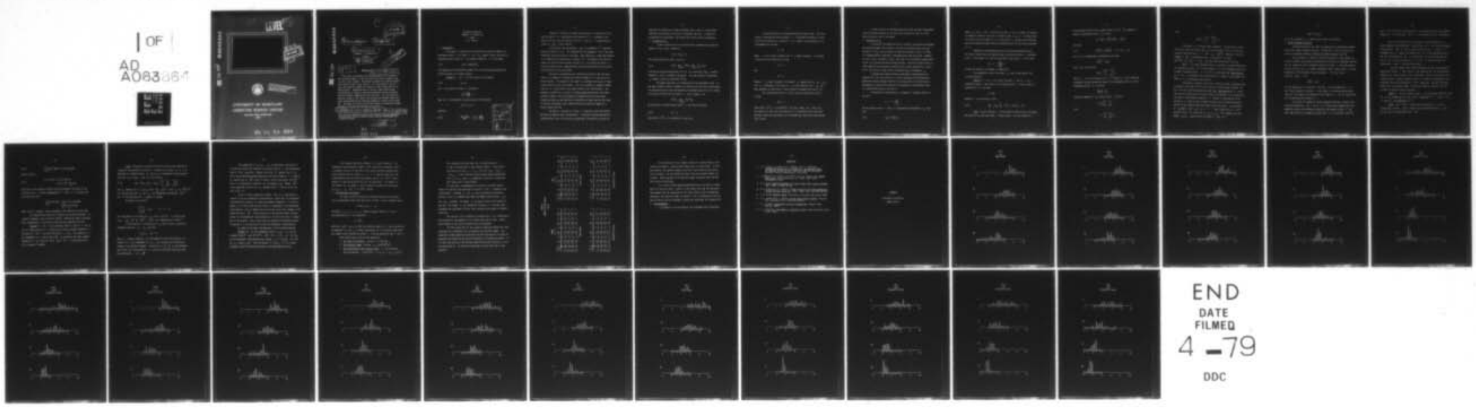
ESTIMATING CONDITION NUMBERS -- AN EMPIRICAL STUDY.(U)

NOV 78 G W STEWART N00014-76-C-0391

UNCLASSIFIED CSC-TR-712 NL

1 OF 1

AD A063864

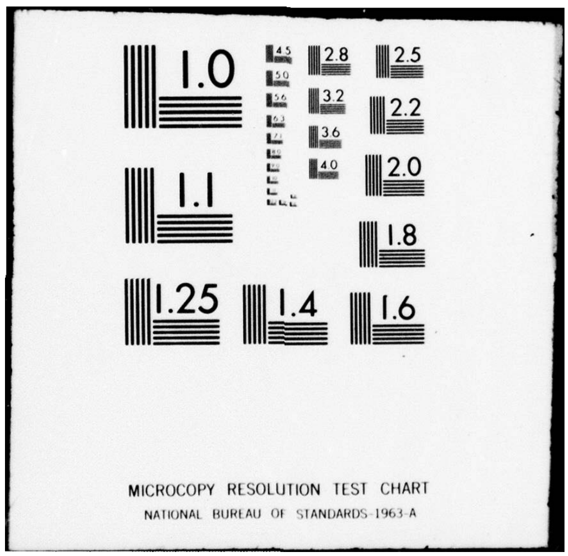


END

DATE FILMED

4-79

DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNIVERSITY OF MARYLAND
COMPUTER SCIENCE CENTER

COLLEGE PARK, MARYLAND
20742

(1)

(9) Technical Report, TR-712   (11) November 1978

(6) Estimating Condition Numbers -- an Empirical Study,

(10) G. W. Stewart*

(12) 37 p.

Abstract (14) CSC-TR-712

This paper investigates two proposed methods for estimating the condition number of a matrix from factorizations commonly used to solve linear systems. One method estimates the condition number with respect to the 1-norm from the LU factorization, and the other the condition number with respect to the 2-norm from the QR factorization. Random matrices of various orders having known distributions of singular values were generated and the estimated condition numbers compared with the true ones. For the classes of matrices tested in this study, the estimators performed rather well, never underestimating the condition number by a factor of more than ten. The paper also gives an efficient method for generating random orthogonal matrices with the Haar distribution.

(15) N00014-76-C-0391, NSF-MSC76-03297

*This work was supported in part by the National Science Foundation and the Office of Naval Research under Contracts MSC76-03297 and N 00014-76-C-0391.

Estimating Condition
Numbers -- an Empirical
Study

G. W. Stewart

1. Introduction.

This paper is concerned with estimating the condition number of a nonsingular matrix A of order n. Let $\|\cdot\|_p$ denote a vector norm and its subordinate matrix norm [7]. The condition number of A is the number

$$(1.1)\qquad \kappa_p(A) = \|A\|_p \|A^{-1}\|_p .$$

The importance of the condition number lies in the following theorem (see [7,8] for proofs of a slight variant).

Theorem 1.1. Let x be the solution of the system

$$(1.2)\qquad Ax = b.$$

Let E be a matrix of order n satisfying

$$\kappa_p(E) \frac{\|E\|_p}{\|A\|_p} < 1.$$

Then $A+E$ is nonsingular, and the solution of the system

$$(A + E)\,\bar{x} = b$$

satisfies

$$(1.3)\qquad \frac{\|x - \bar{x}\|_p}{\|\bar{x}\|_p} \le \kappa_p(A) \frac{\|E\|_p}{\|A\|_p} .$$

Theorem 1.1 allows us to assess the accuracy of a solution of (1.2) when the matrix A is perturbed by the addition of E. Informally, the bound (1.3) says that the relative error in A due to E is magnified by a factor of $\kappa_p(A)$ in the solution.

In principal, one can estimate $\kappa_p(A)$ by computing A^{-1} and estimating the norms in (1.1). This approach has two drawbacks. First, the norms themselves may be difficult to calculate. More important, in many applications it is expensive and unnecessary to compute A^{-1}. For example, the usual procedure for solving linear systems such as (1.2) is to factor the matrix A into the product of two computationally more tractible matrices and use this factorization to solve the system.

One way of circumventing the difficulties raised in the last paragraph is to attempt to estimate the condition number of A from one of its factorizations. The purpose of this paper is to present the results of simulation studies for two such methods. The first estimates a condition number from the LU factorization, the second from the QR factorization. The methods are applied to a variety of randomly generated matrices with known condition numbers, and the estimates are compared with the true values. Speaking broadly, we conclude that the methods are rather good for the class of matrices considered, never underestimating the condition number by a factor greater than ten.

The paper is organized as follows. In Sections 2 and 3 we briefly describe the methods under consideration. In Section 4 we describe how the random matrices used in the study are generated; in particular, we give an

algorithm for generating a random orthogonal matrix that is in some sense uniformly distributed on the space of orthogonal matrices. In Section 5 we describe our experiments and present the results and our conclusions.

2. Estimation of κ_1.

In this section we will be concerned with estimating the condition number for the 1-norm defined by

$$\|x\|_1 = \Sigma_{i=1}^n |x_i|.$$

The subordinate matrix norm is given by

$$(2.1) \qquad \|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|.$$

From the second characterization in (2.1), it is seen that $\|A\|_1$ is easily computed in $O(n^2)$ arithmetic operations. Thus the problem of estimating κ_1 reduces to that of estimating $\|A^{-1}\|_1$.

The method used here is an elaboration of one described in [4]. It has been treated in detail elsewhere [1], and here we give only a brief sketch. The underlying idea is that the maximum in the middle term of (2.1) is likely to be nearly attained for any vector x. Thus to approximate

$$\|A^{-1}\|_1 = \max_{\|u\|_1=1} \|A^{-1}u\|_1,$$

we start with a suitably chosen vector u and solve the system

$$(2.2) \qquad Av = u,$$

after which $\|A^{-1}\|_1$ is estimated by $\|v\|_1 / \|u\|_1$.

Two modifications are incorporated into this basic idea. The first insures a high-quality starting vector u. If Gaussian elimination is used to solve the system (2.2), the matrix A, or rather a row permutation of it, is decomposed in the form

$$A = LU,$$

where L is unit lower triangular and U is upper triangular. The system is then solved by succesively solving

$$Lu' = u$$

and

$$Uv = u'.$$

Because L is lower triangular, the element u_i' depends only on $u_1', \ldots, u_{i-1}'$ and u_i. Consequently, the choice of u_i can be deferred until u_{i-1}' has been computed, at which point it may be chosen to enhance the size of u_i'.

The second modification is to pass once more through the matrix, solving

$$A^T w = v,$$

after which $\|A^{-1}\|_1$ is estimated by $\|w\|_\infty / \|v\|_\infty$, where $\|w\|_\infty = \max \{ w_i \}$. The reasons for this step are detailed in [1]; essentially, the first pass produces a good starting vector for the second pass, where the actual estimation is done.

A slight variant of the method described above has been incorporated into the LINPACK programs for solving linear systems [2], and it is this version that we study in this paper.

3. Estimation of κ_2.

In principal the method outlined in Section 2 can be used to estimate the condition number with respect to any easily evaluated vector norm (an estimate of $\|A\|_p$ can be obtained in much the same way as that of $\|A^{-1}\|_p$, if a computationally convenient characterization does not exist). However, for the 2-norm defined by $\|x\|_2^2 = \Sigma x_i^2$, there is an alternative method based on the QR factorization with column pivoting. Since this factorization is often computed in the course of solving least squares problems (e.g. see [3, 6,7]), we include a study of this alternative way of estimating κ_2.

A computational method based on Householder transformations for computing the QR factorization is closely connected with the method for generating random orthogonal matrices to be described in the next section. Consequently we sketch the properties of Householder transformations here. For computational details see [3,6,7].

A Householder transformation is a symmetric, orthogonal matrix of the form

$$H = I - 2\frac{uu^T}{\|u\|_2^2}.$$

For any nonzero vector a there is a Householder transformation H_a such that

$$H_a a = \mp \|a\|_2^2 e_1, \qquad (3.1)$$

where $e_1 = (1,0,\ldots,0)^T$. In fact we may take $u = e_1 \pm a$, where, for reasons of numerical stability, the sign is chosen to make $\pm a_1 \geq 0$. It follows from (3.1) that the first column of H_a is the vector a scaled to have norm one, while the remaining columns form an orthonormal basis for the space orthogonal to A.

Householder transformations may be stored and manipulated easily. For their storage, they only require space to retain the vector u. Moreover, given u, the product Ha can be calculated for any vector a in the form

$$Ha = a - \frac{2u^T a}{\|u\|_2^2} u,$$

without the necessity of forming H explicitly.

The decomposition we will estimate κ_2 from is described in the following theorem.

Theorem 3.1. Let A be a matrix of order n and let J be a permutation matrix. Then there is an orthogonal matrix Q and an upper triangular matrix R such that

$$A = QRJ.$$

Moreover J can be chosen so that

$$r_{kk}^2 \geq \Sigma_{i=k}^{j} r_{ij}^2 \quad (k=1,\ldots,n; j=k,\ldots,n). \tag{3.2}$$

Proof. We will show how J can be chosen so that (3.2) is satisfied, the proof for an arbitrary fixed J being similar. Let the columns of A

be rearranged so that one of largest 2-norm is first. This amounts to choosing a permutation matrix J_1 so that

$$A_1 = AJ_1^T = (a_1^{(1)}, a_2^{(1)}, \ldots, a_n^{(1)})$$

satisfies

$$\| a_1^{(1)} \|_2 \geq \| a_j^{(1)} \|_2 \qquad (j = 2,3,\ldots,n). \tag{3.3}$$

Let H_1 be a Householder transformation such that

$$H_1 a_1^{(1)} = r_{11} e_1.$$

Then $H_1 A_1$ has the form

$$H_1 A_1 = \begin{pmatrix} r_{11} & r_1^T \\ 0 & A_2 \end{pmatrix},$$

where r_1 is an (n-1)-vector and A is of order n-1. By an induction hypothesis, we may assume that there is an orthogonal matrix Q_2 and a permutation matrix J_2 such that

$$Q_2^T A_2 J_2^T = R_2,$$

where the elements of R_2 satisfy (3.2). If we set

$$Q = Q_1 \begin{pmatrix} 1 & 0 \\ 0 & Q_2 \end{pmatrix}$$

and

$$J = \begin{pmatrix} 1 & 0 \\ 0 & J_2 \end{pmatrix} J_1,$$

then

$$Q^T A J^T = \begin{pmatrix} r_{11} & r_1^T J_2 \\ 0 & R_2 \end{pmatrix} \equiv R$$

The matrix R is clearly upper triangular. To show that it satisfies (3.2), note first that since R_2 also satisfies (3.2), it is only necessary to treat the case $k = 1$. Now $r_{11} = \|a_1^{(1)}\|$. Moreover if the j-th element of $r_1^T J_2$ corresponds to the ℓ-th column of A_1, then $\|a_\ell^{(1)}\|_2^2 = r_{1j}^2 + r_{2j}^2 + \ldots + r_{jj}^2$. The result now follows from (3.3). ∎

When J is fixed, the above proof is essentially a specification for an algorithm to factor AJ^T into the product QR of an orthogonal matrix and an upper triangular matrix. If $J=I$, this is called <u>the QR factorization of</u> A. For A nonsingular, this factorization is unique if the diagonal elements of R are required to be positive.

The process, sketched in the proof, by which columns of maximal norms are moved to the front of the current matrix is called column pivoting. We will call the resulting decomposition <u>the QRJ factorization of</u> A. Unless there are ties among norms of columns during the reduction, it is unique up to the signs of the diagonal elements of R. Because of the condition (3.2), the matrix R tends to show a graded structure with the elements becoming smaller as one procedes toward the lower right hand corner. In particular, if $\text{rank}(A)=r$ then $r_{ij}=0$ for $j \geq i > r$. Near singularity of A will manifest itself in a large ratio of r_{11} to r_{nn}. This suggests that the number r_{11}/r_{nn} may provide an estimate of $\kappa_2(A)$; i.e.,

$$\kappa_2(A) \cong r_{11}/r_{nn}.$$

It is this estimate of κ_2 *that we shall study in this paper.*

4. Random Orthogonal Matrices.

The matrices used in our study are generated by transforming diagonal matrices by random orthogonal matrices. Accordingly, in this section we describe an economical method for generating random orthogonal matrices.

Just as there is a natural uniform distribution for real numbers, so is there a natural distribution associated with the space of orthogonal matrices. Let $\mathcal{O}_n$ denote the set of orthogonal matrices of order n. With the spectral norm as a metric, $\mathcal{O}_n$ is a compact topological group. Hence there is a unique normalized, left-translation invariant measure μ for $\mathcal{O}_n$; that is, μ satisfies $\mu(\mathcal{O}_n) = 1$ and

$$\mu(H\mathcal{X}) = \mu(\mathcal{X})$$

for any measurable $\mathcal{X} \subset \mathcal{O}_n$ and any $H \in \mathcal{O}_n$ [5]. It can be shown that μ is also right translation invariant. In the sequel all functions and distributions will be assumed to be measurable with respect to μ and all integrations will be performed with respect to μ.

Our method for generating random orthogonal matrices is based on the following observation. Let X be an $n \times n$ matrix whose elements are independently normally distributed with mean zero and common variance σ^2 (we shall abbreviate this statement by saying that X is distributed $NI(0,\sigma^2)$).

Let $X = QR$ be the QR factorization of X, normalized so that the diagonal elements of R are positive. Then Q is distributed over $\mathcal{O}_n$ according to μ.

This fact follows from the invariance of the distribution of X under orthogonal transformations. Specifically, let H be a fixed orthogonal transformation. Then HX has the same distribution as X. Since HQ is orthogonal and $HX = (HQ)R$, the random matrix HQ is the orthogonal part of the QR factorization of HX. It follows that HQ has the same distribution as Q; i.e. the distribution of Q on $\mathcal{O}_n$ is invariant under left translations. Hence Q has the distribution μ.

All this suggest that random orthogonal matrices be generated by *generating a matrix* X *from* $NI(0,\sigma^2)$ *and computing its* QR *factorization.* However this process requires $O(n^3)$ operations. We shall show that we can instead express a random orthogonal matrix as a product of Householder transformations that can be generated in $O(n^2)$ operations. We begin our development with an elementary lemma, in which we extend our notation $NI(0,\sigma^2)$ to refer to vectors as well as matrices.

Lemma 4.1. Let the random n-vector x be distributed $NI(0,\sigma^2)$ and let H be a random matrix on $\mathcal{O}_n$ that is independent of x. Then $y = Hx$ is $NI(0,\sigma^2)$ and independent of H.

Proof. Let f and g denote the distributions of x and H respectively, and let $h(y,H)$ denote the joint distribution of y and H. Let $\mathcal{Y}$ and $\mathcal{H}$ be measurable sets. Then

$$\int_{\mathcal{H}}\int_{\mathcal{Y}} h(y,H)\ dydH = \Pr\left[(y,H) \in \mathcal{Y}\times\mathcal{H}\right]. \tag{4.1}$$

Now for fixed H

$$\begin{aligned}\Pr[y \in \mathcal{Y} \mid H] &= \Pr[x \in H^T\mathcal{Y} \mid H] \\ &= \Pr[x \in \mathcal{Y}] = \int_{\mathcal{Y}} f(x)dx.\end{aligned} \tag{4.2}$$

Then next to last equality follows from the orthogonal invariance of the distrib-tion of x and the independence of x from H. It follows from (4.1) and (4.2) that

$$\begin{aligned}\int_{\mathcal{H}}\int_{\mathcal{Y}} h(y,H)\ dydH &= \int_{\mathcal{H}} g(H)\ \Pr\left[y \in \mathcal{Y} \mid H\right]dH \\ &= \int_{\mathcal{H}} g(H) \int_{\mathcal{Y}} f(x)\ dxdH.\end{aligned}$$

Hence $h(y,H) = f(y)g(H)$ almost everywhere, which implies the result. ■

The computational algorithm is based on the following theorem. Its complete statement is due to Garrett Birkhof, although the author discovered parts independently in the course of deriving the algorithm to follow.

Theorem 4.2. Let X be distributed $NI(0,\sigma^2)$ and let $X = QR$ be the QR factorization of X normalized so that the diagonal elements of R are positive. Then Q has the distribution μ on $\mathcal{O}_n$. The elements of R are independent of Q and of each other. In particular the super diagonal elements of R are $NI(0,\sigma^2)$, while r_{ii}^2/σ^2 has χ^2 distribution with $n-i+1$ degrees of freedom.

<u>Proof</u>. We consider a variant of the first step of the reduction to triangular form sketched in Section 3. From the first column x_1 of X, we determine an orthogonal matrix H_1 (which is ± a Householder transformation) such that $H_1 x_1 = \|x_1\| e_1$. Then $H_1 X$ has the form

$$(4.3) \qquad H_1 X = (H_1 x_1, H_1 x_2, \ldots H_1 x_n) = \begin{pmatrix} r_{11} & r_{12} \cdots & r_{1n} \\ 0 & x_2^{(2)} \cdots & x_n^{(3)} \end{pmatrix}.$$

Now H_1 depends only on $x_1/\|x_1\|$, while $r_{11}^2 = \|x_1\|^2$. Since x_1, is $NI(0,\sigma^2 I)$, it is easily seen that H_1 and r_{11} are independent and moreover r_{11}^2/σ^2 has χ^2 distribution with n degrees of freedom.

By Lemma 4.1 the vectors

$$\begin{pmatrix} r_{1j} \\ x_j^{(2)} \end{pmatrix} \equiv H_1 x_j \qquad (j = 2,\ldots,n)$$

are independent of each other and H_1 and are $NI(0,\sigma^2)$. It follows that $(r_{12},\ldots,r_{1n})$ and $X_2 = (x_2^{(2)},\ldots,x_n^{(2)})$ are independent and $NI(0,\sigma^2)$.

If the reduction is continued with X_2, there results a sequence of orthogonal matrices, $H_1,\ldots,H_n$ such that

$$H_n H_{n-1} \cdots H_1 X = R,$$

where R is upper triangular. By the arguments of the last paragraph, the elements of R are independent of $H_1,\ldots, H_n$ and have the distributions claimed in the theorem statement. The matrix $Q = H_1^T \ldots H_n^T$ is the orthogonal part of the QR factorization of X and by the discussion above must have the distribution μ on $\mathcal{O}_n$. ■

The computation of $H_1, H_2, \ldots, H_n$ as described in the proof of the theorem involves the reduction of the entire matrix X and consequently entails $O(n^3)$ operations. However, note that H_1 depends only on x_1, which can be generated without generating the other columns of X. Likewise, H_2 depends only on $x_2^{(2)}$, which is simply a vector distributed $NI(0,\sigma^2)$. Thus it is unnecessary to generate x_2 and compute $H_1 x_2$. Rather, $x_2^{(2)}$ can be generated directly, and H_2 computed from it -- and so on for the rest of the H's.

There is a minor annoyance to remove. The H's in the proof of Theorem 4.2 are not Householder transformations; rather they are Householder transformations scaled by ±1 to make the diagonal elements of R positive. However, it is easily seen that the factor is independent of the Householder transformation, since it depends only on the sign of the first element in the generating vector $x_j^{(j)}$. Hence the matrix of the theorem differs from the product of the Householder transformations by a factor of ±1 that is independent of the product. Such a factor does not change the distribution; hence in generating Q we need work only with the Householder transformations.

We summarize the above considerations in the following theorem.

<u>Theorem</u> 4.3. Let the independent vectors $x_1, x_2, \ldots, x_n$ be distributed $NI(0,\sigma^2)$ over $\mathbb{R}^n, \mathbb{R}^{n-1}, \ldots, \mathbb{R}^1$. For $j = 1,2,\ldots,n$, let $\bar{H}_{x_j}$ be the Householder transformation that reduces x_j to $\mp\|x_j\|e_1$ cf. (3.1). Let $H_j = \mathrm{diag}(I_{j-1}, \bar{H}_j)$. Then the product $Q = H_1 H_2 \cdots H_n$ is a random orthogonal matrix distributed according to the Haar measure over $\mathcal{O}_n$.

The procedure implied by Theorem 4.1 is quite economical. The Householder transformations require $O(n^2)$ operations to generate, and it is probable that most of the work will be spent generating pseudo-random normal deviates for the vectors x_j. If Q is required explicitly, the transformations must be multiplied together, which is an $O(n^3)$ process. However, in many applications this will be unnecessary. For example, if the product Qx is needed, it can be formed by successively multiplying $H_n, H_{n-1}, \ldots, H_1$ into x, an $O(n^2)$ process.

5. The Experiment and Results.

In order to test the condition estimaters described in Sections 2 and 3, we generated random test matrices as follows. Given a diagonal matrix

$$\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

satisfying $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$, random orthogonal matrices U and V were generated and A was computed as

$$A = U \Sigma V^T.$$

Note that $\kappa_2(A) = \sigma_1/\sigma_n$, so that the condition number of A can be controlled by adjusting σ_1 and σ_n. Further replications of A having the same condition number can be obtained by holding Σ fixed and generating new U and V.

Three factors were varied in the experiment.

1. The order of the matrix. We took $n = 5, 10, 25, 50$.
2. The condition number. We took $\kappa_2 = 10, 10^2, 10^4, 10^6$.
3. The distribution of the singular values σ_i. We considered two distributions. In the first $1 = \sigma_1 = \sigma_2 = \ldots = \sigma_{n-1} > \sigma_n = \kappa_2^{-1}$.

This represents the case where the ill-conditioning of A is due to a sharp break in the singular values. In the second distribution we took $\sigma_1 = 1$, $\sigma_n = \kappa_2^{-1}$, and $\sigma_2/\sigma_1 = \sigma_3/\sigma_2 = \ldots = \sigma_n/\sigma_{n-1}$. In this case the singular values decay exponentially from σ_1 to σ_p, and it is impossible to point to a single one that causes the ill-conditioning.

For each case, corresponding to a choice of the three factors, twenty-five matrices were generated by varying U and V randomly as described above. Estimates $\hat{\kappa}_1$ and $\hat{\kappa}_2$ were calculated as described in Sections 2 and 3 via LINPACK codes SGECO and SQRDC, and the ratios $\hat{\kappa}_1/\kappa_1$ and $\hat{\kappa}_2/\kappa_2$ recorded. The number κ_2 was given as one of the factors of the case. The number κ_1 was computed by inverting A. *Care was taken* throughout the experiment to insure that rounding error played an insignificant part.

The raw data is too voluminous to present here. It is summarized in the histograms in the appendix to this paper (the scale is x10). Table 1 gives medians of the ratios estimated from the histograms.

The table shows that for the classes of matrices treated here, both techniques do a reasonable job of estimating the condition number. The algorithm in SGECO generally outperforms the QRJ factorization, as might be expected since SGECO was specifically designed as a condition estimater. In all cases the quality of the estimate deterioriates with increasing n, but *not drastically so. In the entire experiment no* ratio less than 0.1 was observed.

Table 1

Median of the Ratios

$\hat{\kappa}_p/\kappa_p$

SGECO

$p = 1$

$n \backslash \kappa_2$	10	10^2	10^4	10^6
5	.74	.73	.73	.75
10	.64	.65	.65	.64
25	.46	.56	.56	.56
50	.44	.50	.50	.50

sharp break

$n \backslash \kappa_2$	10	10^2	10^4	10^6
5	.63	.67	.70	.75
10	.48	.52	.61	.60
25	.32	.35	.39	.48
50	.23	.27	.32	.34

exponential decay

QRJ

$p = 2$

$n \backslash \kappa_2$	10	10^2	10^4	10^6
5	.70	.71	.78	.75
10	.61	.63	.62	.60
25	.44	.44	.44	.46
50	.37	.36	.36	.36

sharp break

$n \backslash \kappa_2$	10	10^2	10^4	10^6
5	.62	.54	.51	.55
10	.50	.40	.34	.34
25	.41	.30	.24	.23
50	.35	.24	.17	.15

exponential decay

The distribution of the singular values has a marked effect on the condition estimaters. Both do better when there is a sharp break. For this distribution, the condition number itself has little effect on the quality of the estimate. For the exponential decay, the two estimaters behave differently. SGECO improves as the condition number increases, while the QRJ factorization deteriorates.

It is risky to make sweeping generalizations on the basis of experiments as limited as these. However, to the extent that they are applicable they suggest that the techniques will be useful when an order of magnitude estimate of the condition number is desired. This is reinforced by the fact that to date no one has succeeded in making the algorithms fail dramatically.

6. Acknowledgement.

I am indebted to William Schwarz, who programmed these experiments.

References

1. A. K. Cline, C. B. Moler, G. W. Stewart, and J. H. Wilkinson, LINPACK working note #7; an estimate for the condition number of a matrix, Argonne National Laboratory, Applied Mathematics Division, Technical Memorandum No. 310 (1977).

2. J. Dongarra, J. R. Bunch, C. B. Moler, and G. W. Stewart, The LINPACK Users Guide, to be published by Soc. Indust. Appl. Math., Philadelphia (1979).

3. G. H. Golub, Numerical methods for solving linear least squares problems, Numer. Math. 7 (1965) 206-216.

4. W. B. Gragg and G. W. Stewart, A stable variant of the secant method for solving nonlinear equations, SIAM J. Numer. Anal. 13 (1976) 889-903.

5. P. R. Halmos, Measure Theory, Van Nostrand, Princeton, New Jersey (1950).

6. C. L. Lawson and R. J. Hanson, Solving Least Squares Problems, Prentice-Hall, Englewood Cliffs, New Jersey (1974).

7. G. W. Stewart, Introduction to Matrix Computations, Academic Press, New York (1974).

8. J. H. Wilkinson, The Algebraic Eigenvalue Problem, Oxford University Press, London (1965).

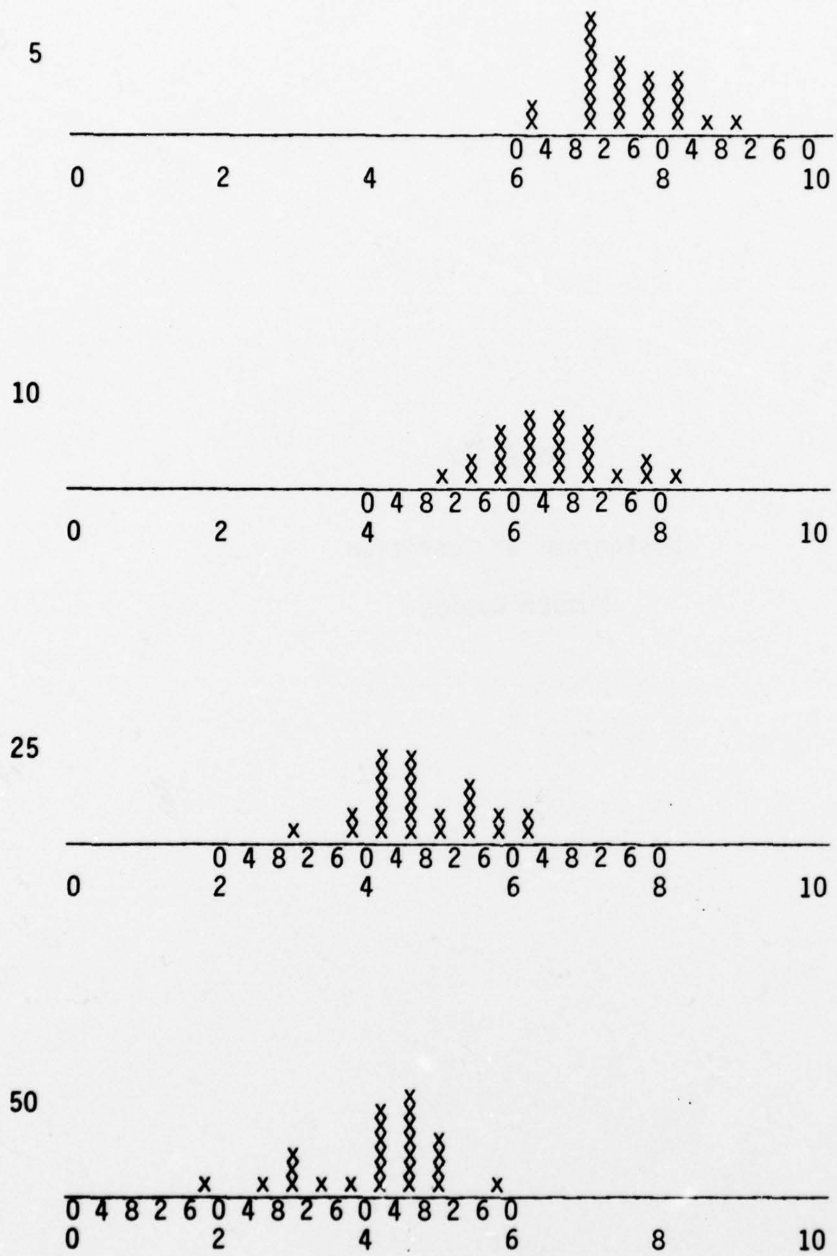
APPENDIX

Histograms of Condition

Number Ratios

SGECO

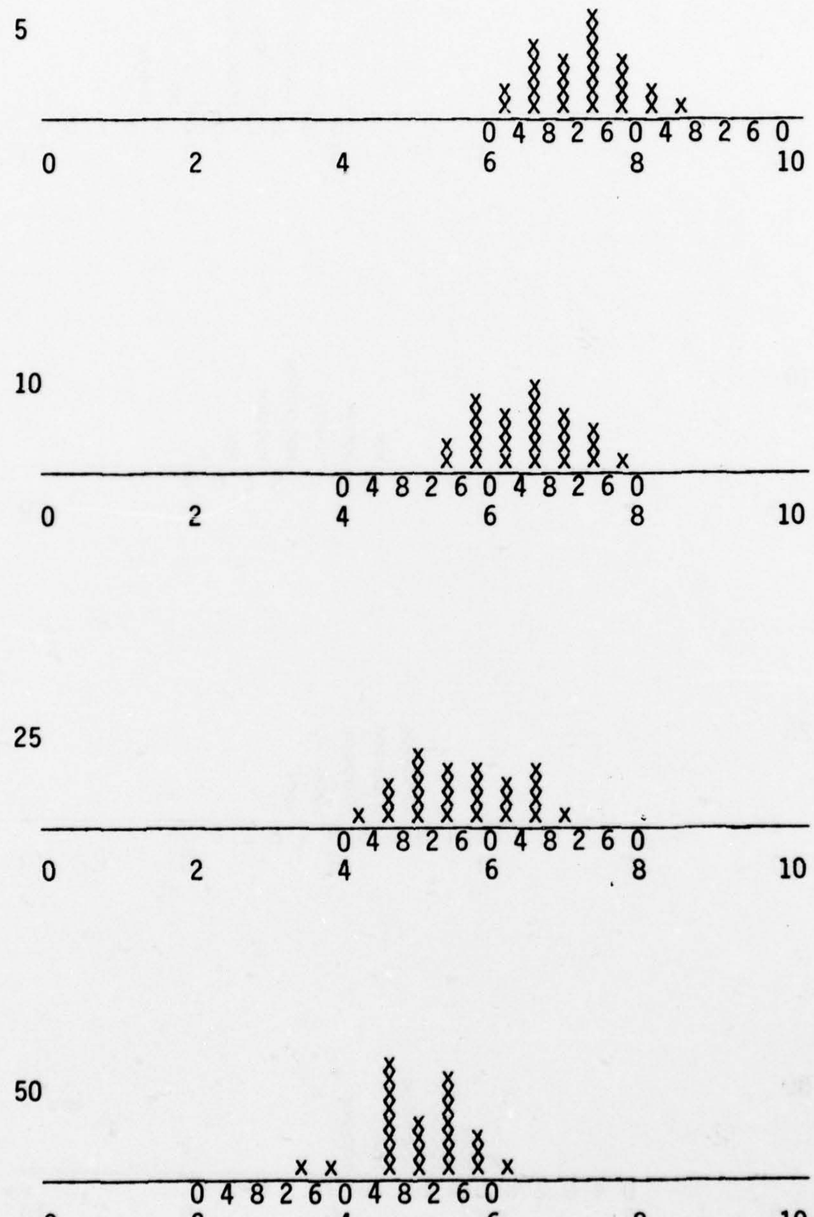
$\kappa=10$
Sharp Break



SGECO

$\kappa=10^2$

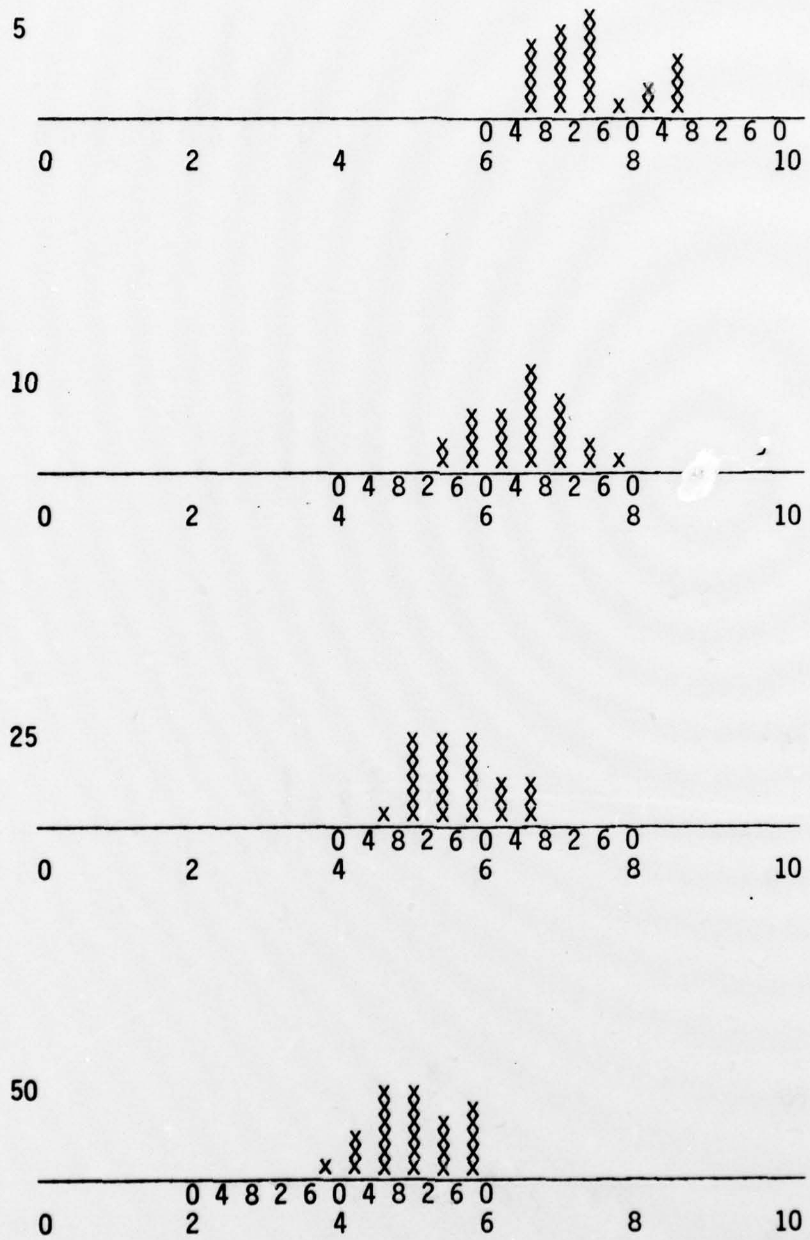
Sharp Break



SGECO

$\kappa=10^4$

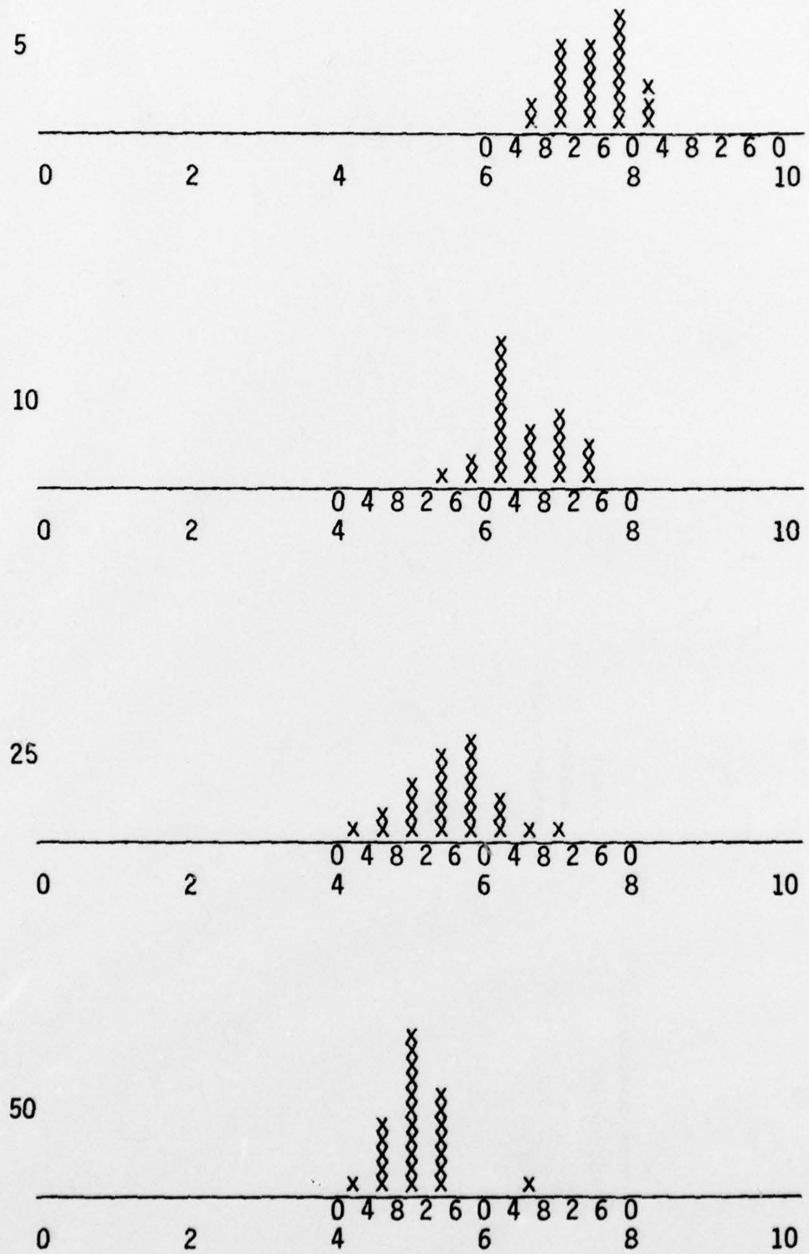
Sharp Break



SGECO

$\kappa=10^6$

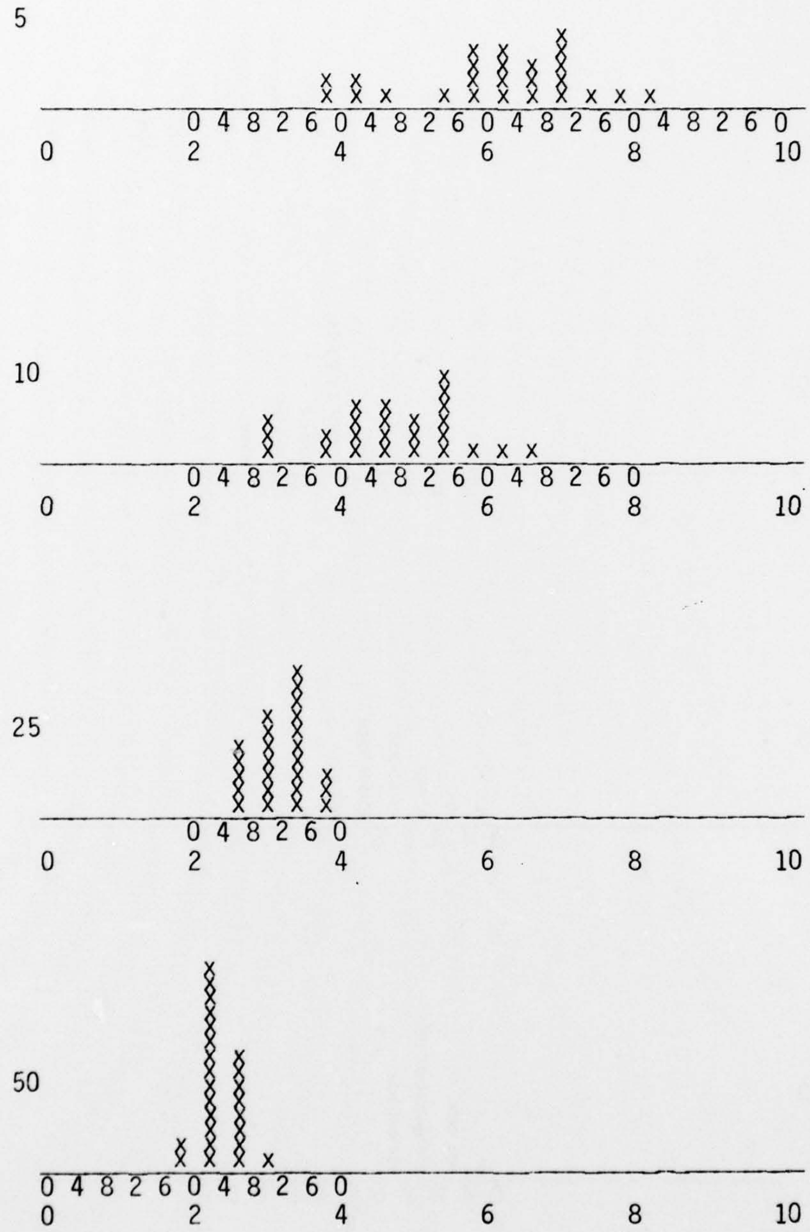
Sharp Break



SGECO

$\kappa=10$

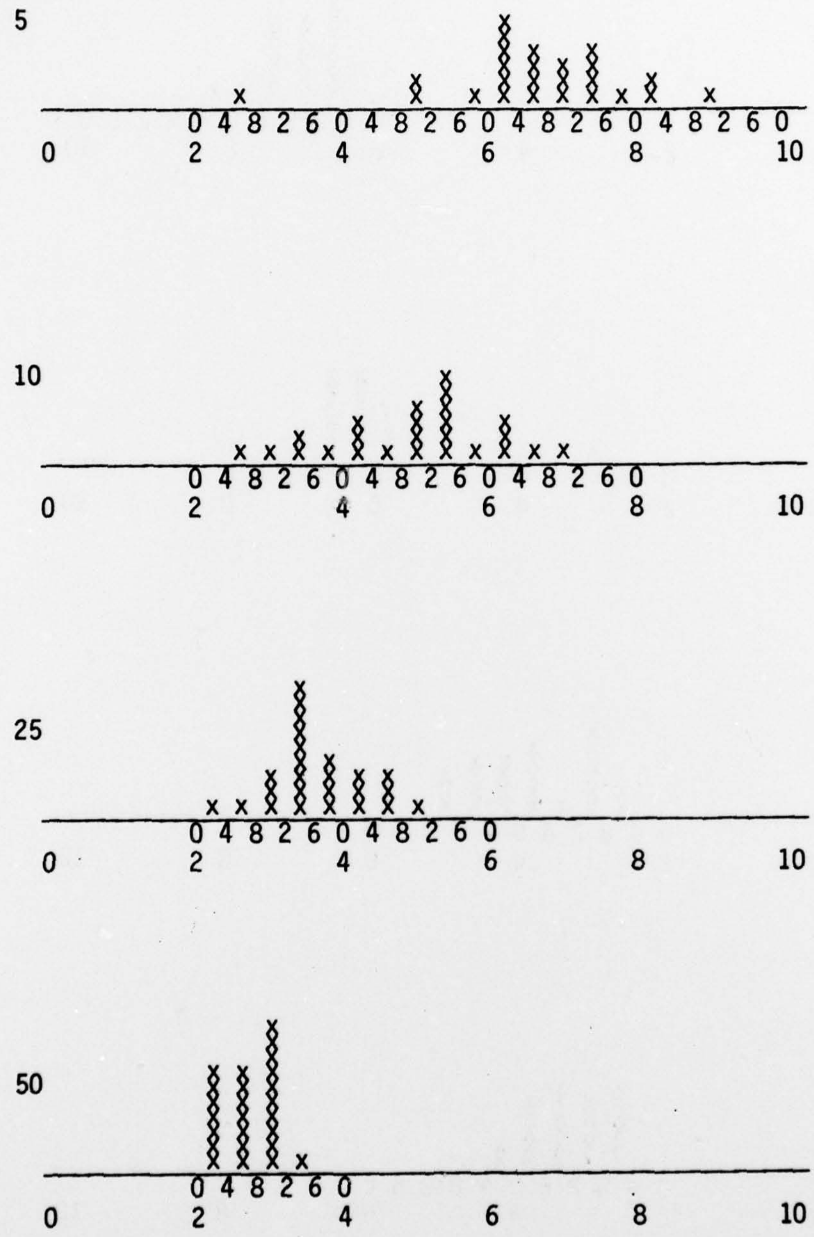
Exponential Decay



SGECO

$\kappa=10^2$

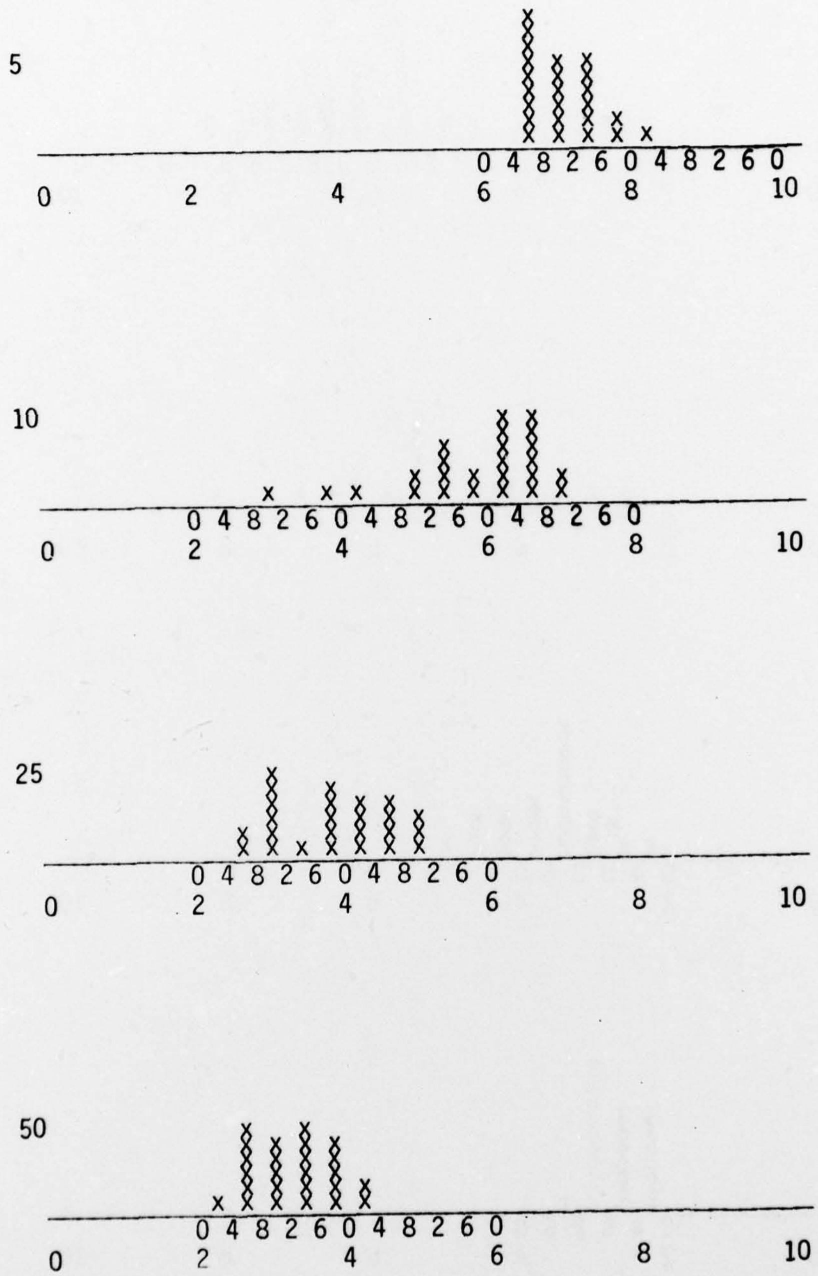
Exponential Decay



SGECO

$\kappa=10^4$

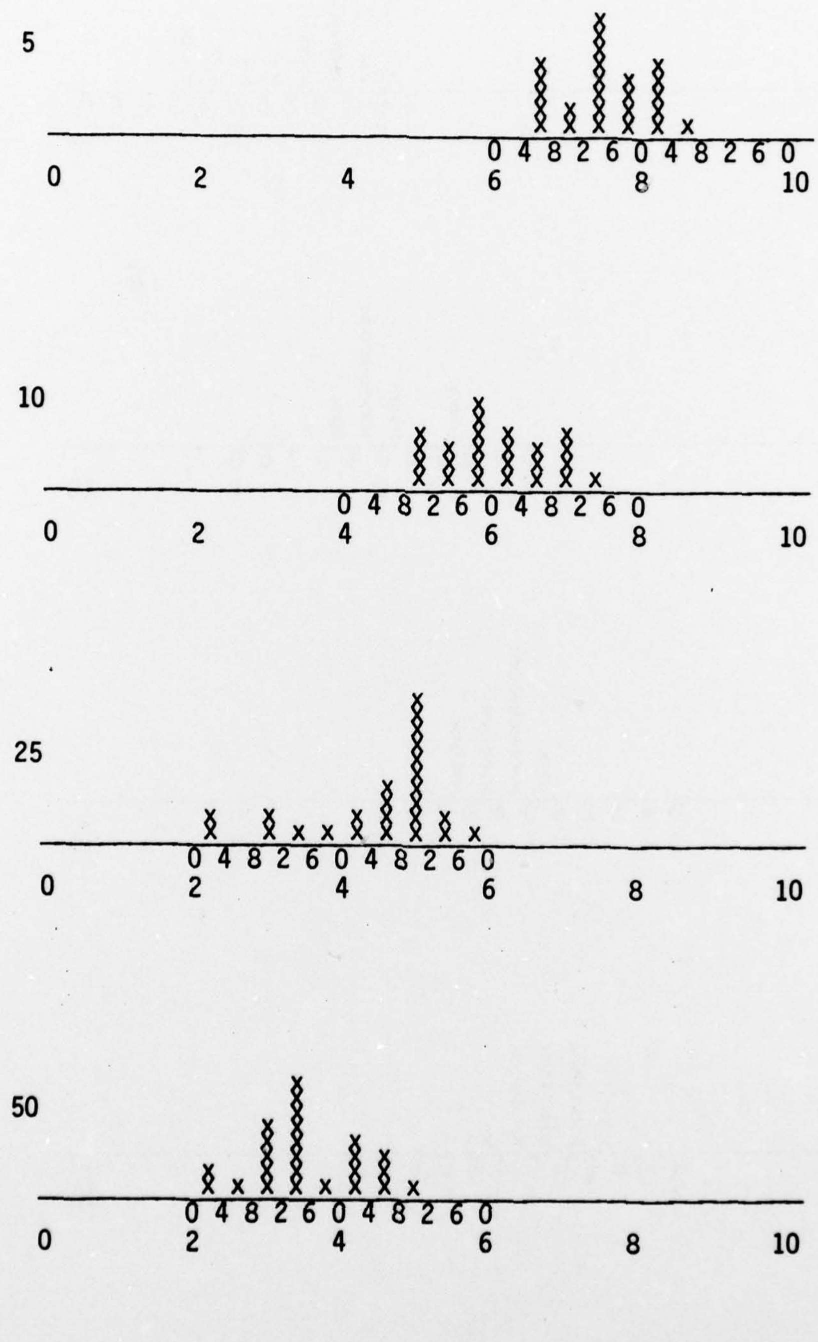
Exponential Decay



SGECO

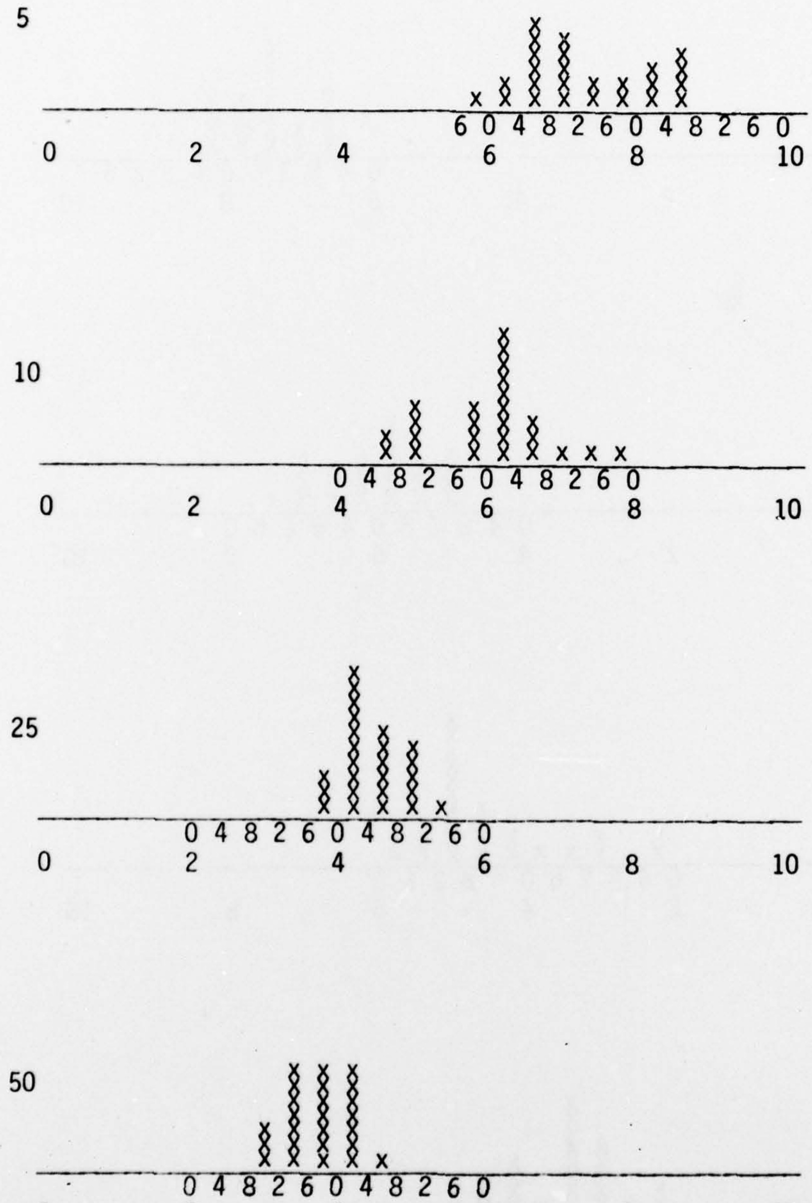
$\kappa = 10^6$

Exponential Decay



QRJ

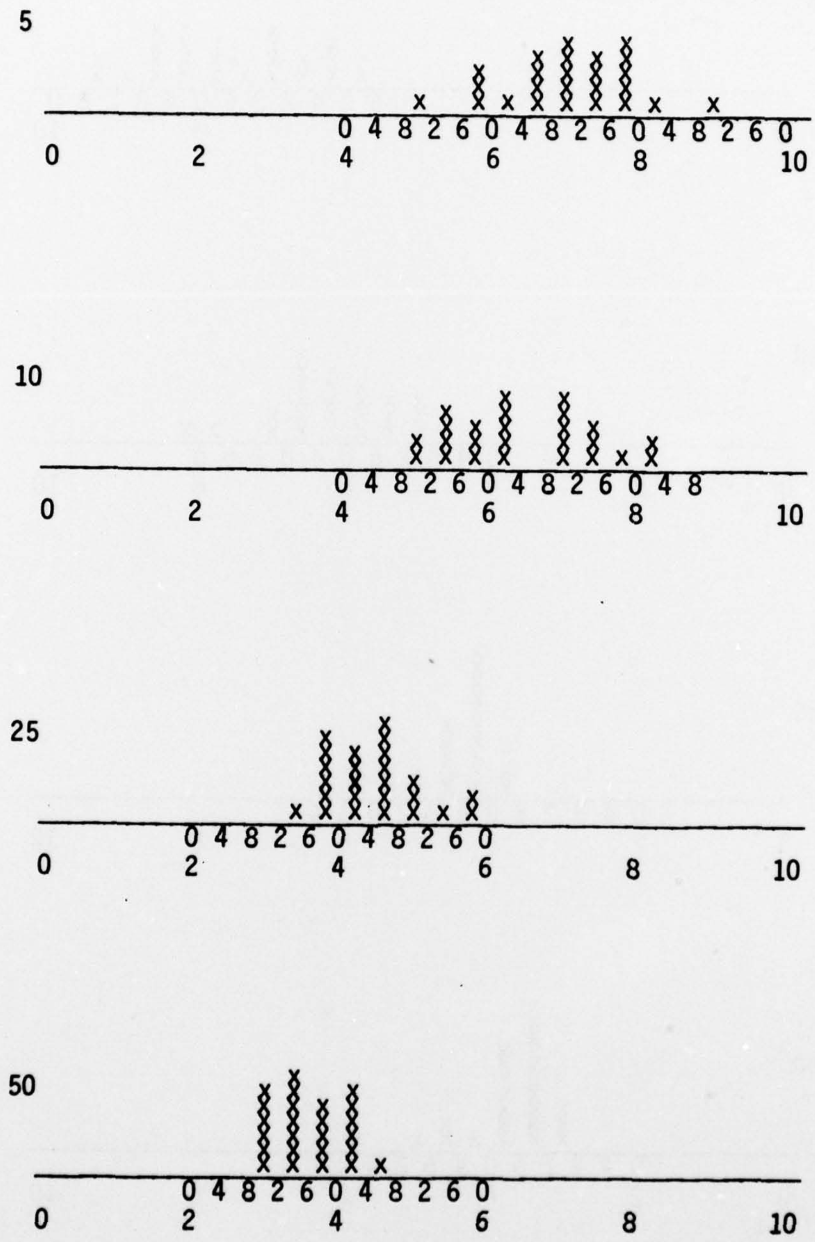
$\kappa = 10$
Sharp Break



QRJ

$\kappa=10^2$

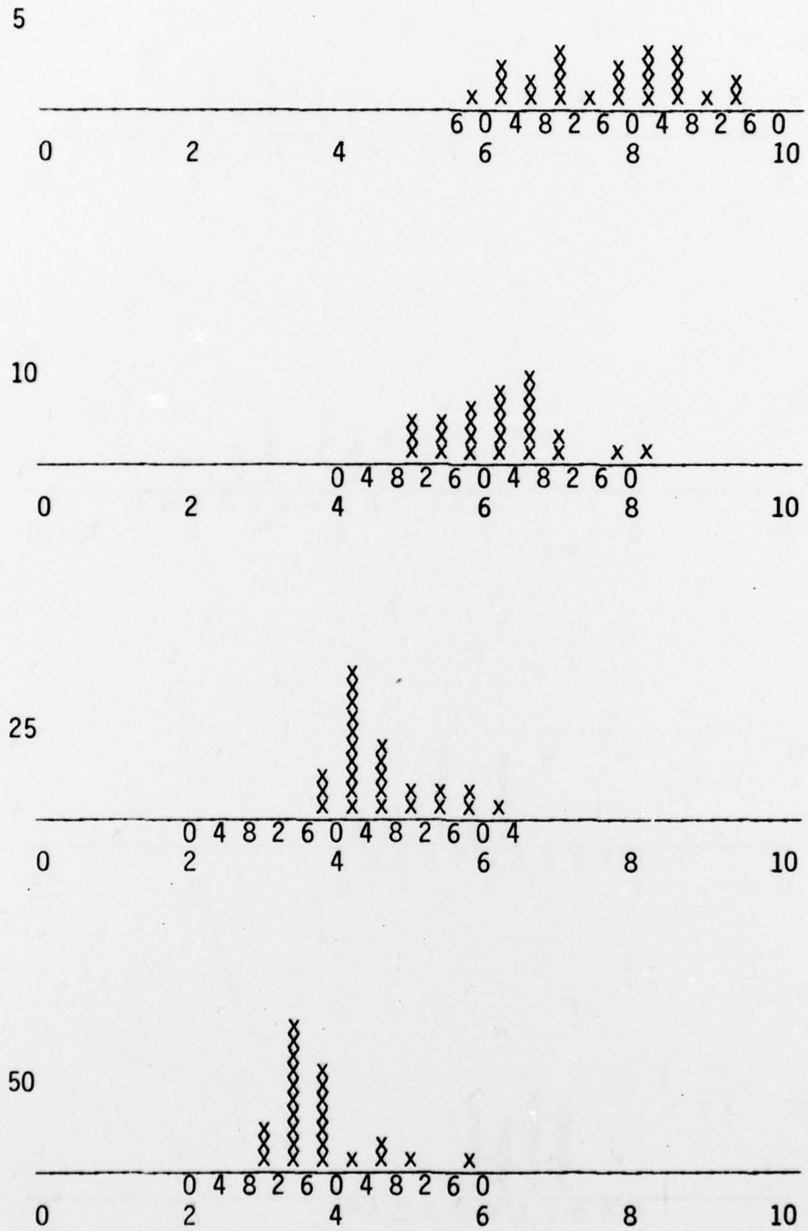
Sharp Break



QRJ

$\kappa=10^4$

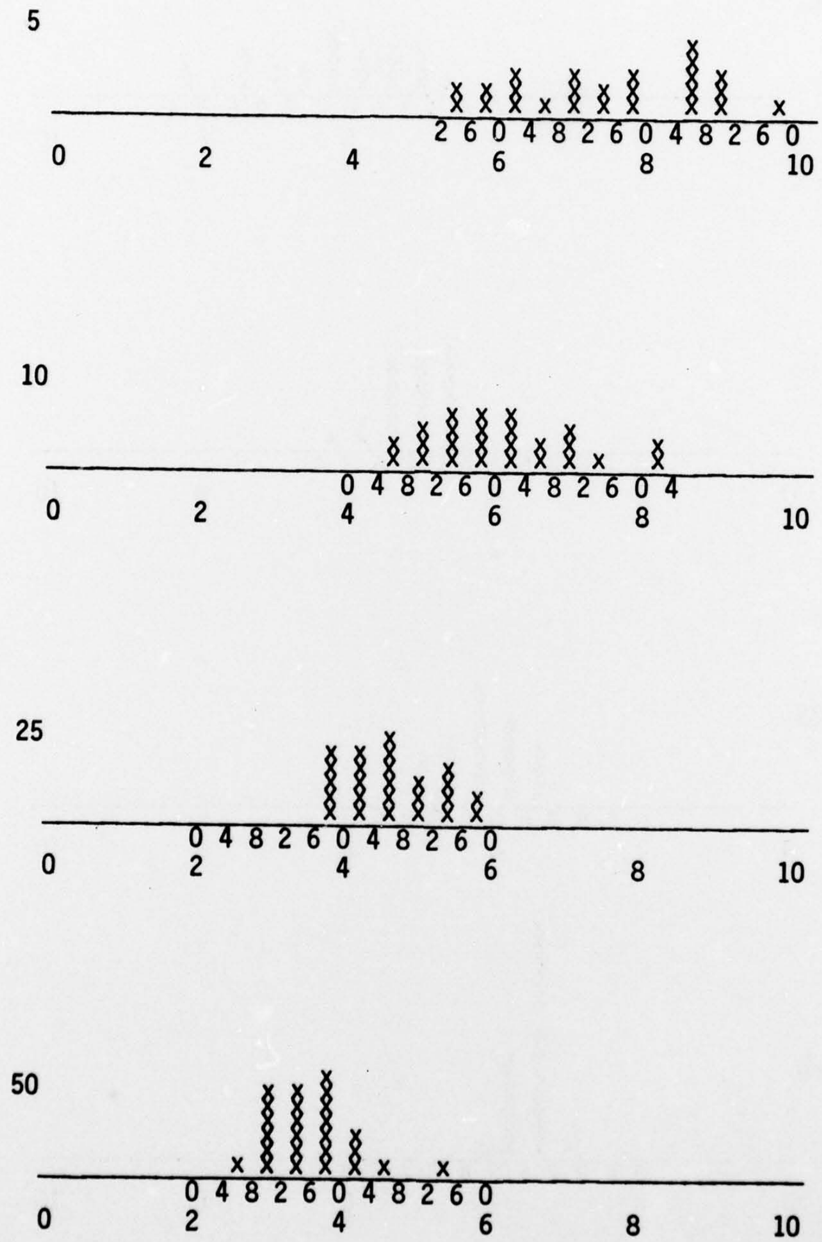
Sharp Break



QRJ

$\kappa=10^6$

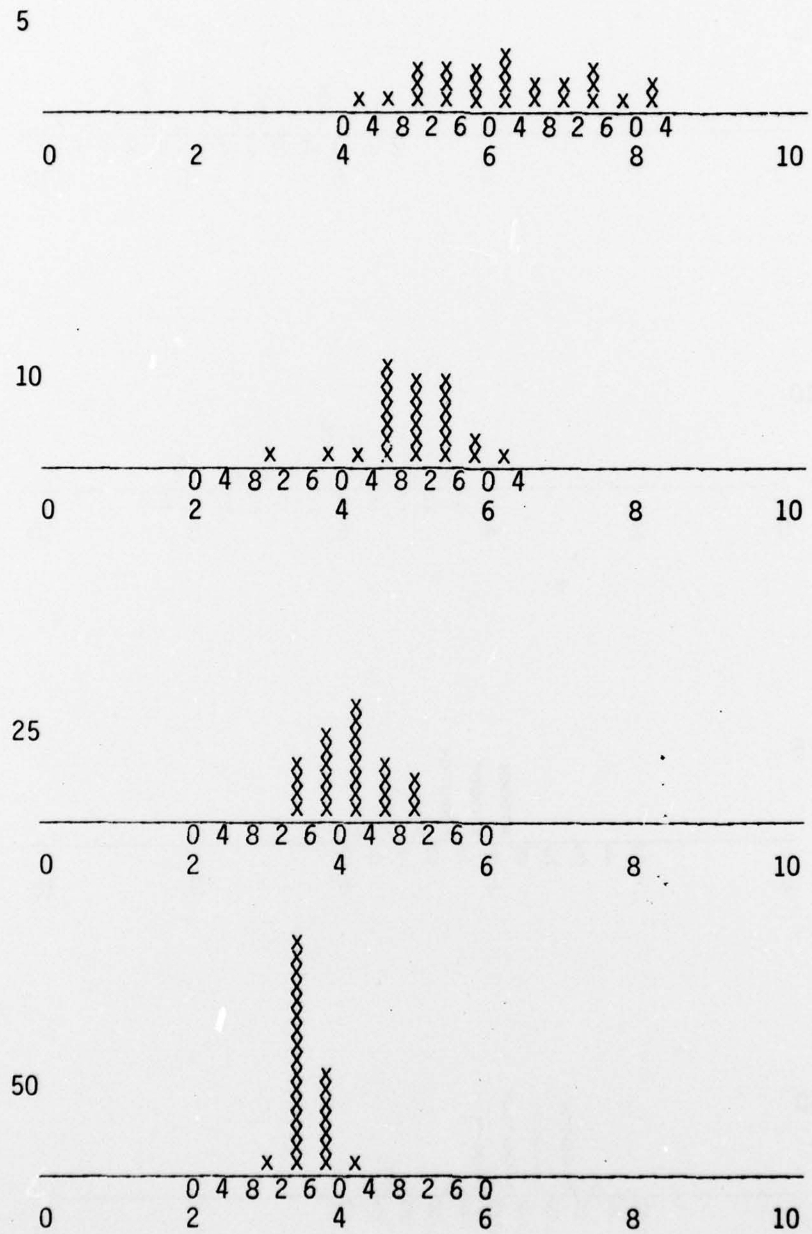
Sharp Break



QRJ

$\kappa=10$

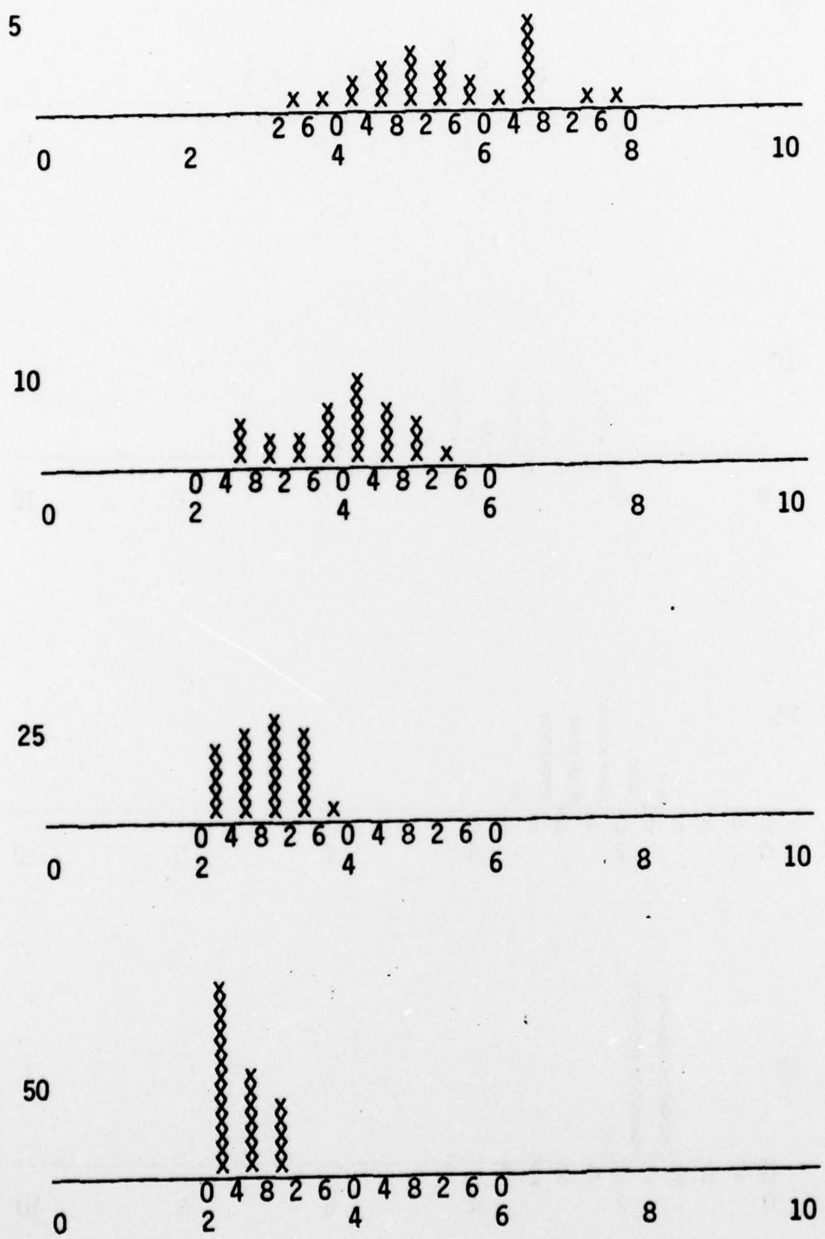
Exponential Decay



QRJ

$\kappa=10^2$

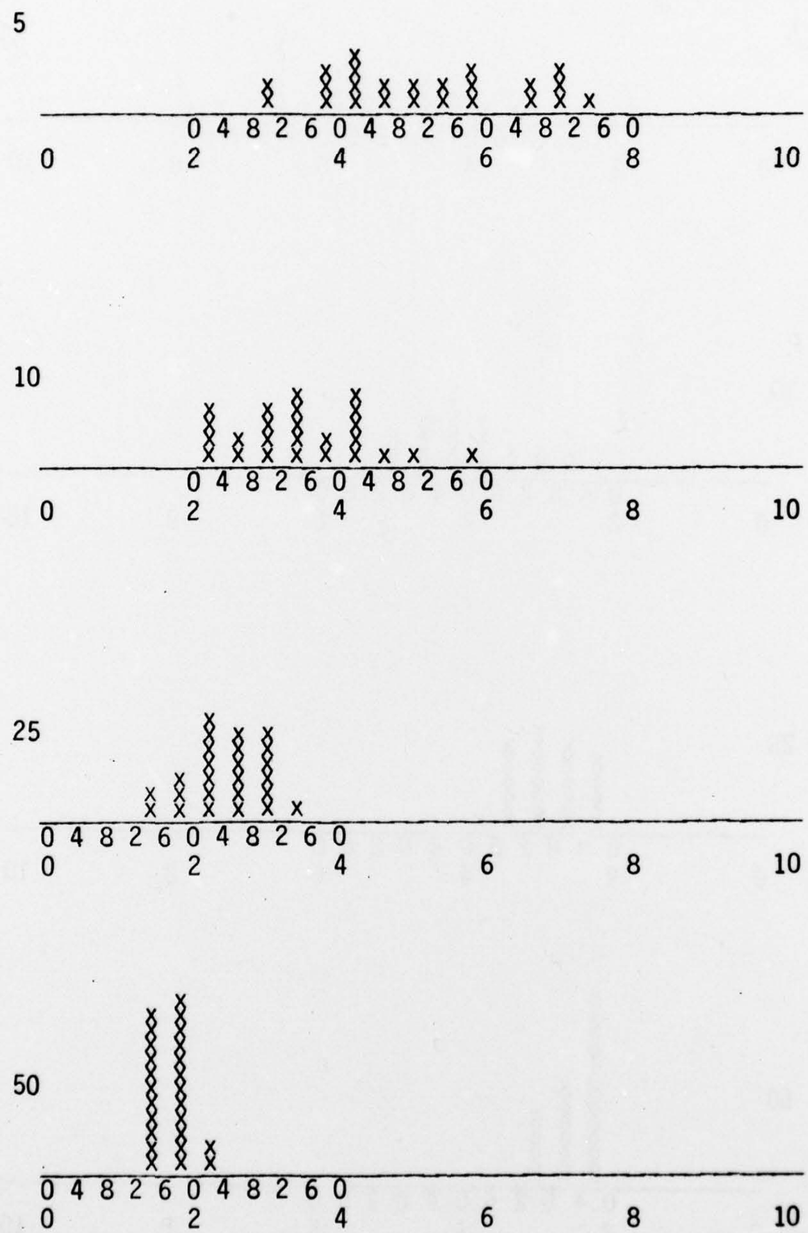
Exponential Decay



QRJ

$\kappa=10^4$

Exponential Decay



QRJ

$\kappa=10^6$

Exponential Decay